Partition of Random Items: Tradeoff between Binning Utility, Meta Information Leakage and Hypotheses Distinguishability

Farhang Bayat, Shuangqing Wei
Department of Electrical and Computer Engineering, Louisiana State University, USA

***Abstract*—In this paper, we propose a novel formulation to understand the tradeoff between channel utility, meta information leakage and hypotheses distinguishability. As an example, such problems could emerge when online users aim to hide some level of information about themselves while revealing enough to be distinguishable from others.**

Under the framework, we formulate two constrained optimization problems. In these problems, the goal is to maximize binning utility while restraining a certain level of information leakage and to maximize the measure of distinguishability between two hypotheses while restraining a certain level of information leakage in the former and latter cases respectively. In both cases, we accomplish this goal by properly dividing a set of M random items into N bins. Such problems (where optimization is carried out over a diverse but yet dependent series of sets) are formally known as multi-agent multi-variate optimization problems.

Both of the constrained optimization problems are NP hard by nature. We then propose suboptimal solutions to exploit potential sub-modular structures of the problems based upon sufficient conditions found on the payoff functions in the underlying objective functions. Further numerical results are presented to demonstrate the effectiveness of the proposed sub-modular function based algorithms.

***Index Terms*—partition, information leakage, privacy, mutual information, hypotheses distinguishability, submodularity**

I. Introduction

In our daily use of internet there are two very specific issues of privacy. The first issue rises when we aim to hide as much vital information about ourselves while still enjoying the service of web -which is always prone to eavesdrops. As an example, we hope to use the Google search engine without a third party studying our search history and learning about us. For future references, we refer to this class of problems as Utility versus Privacy (UvP).

The second scenario rises when we wish to be distinguished from other users but not reveal all the information we share on the web. As an example, when we surf Amazon, based upon our previous searches over the website, we are offered new suggestions we may enjoy. However, we hope to keep our privacy intact to some level at the same time. If we search for a specific political ideology, we do not like to be suggested a book on such an idea. If we were, we would feel uneasy about how much private information about us has leaked. This class of problems will be referred to as Utility versus Privacy and Distinguishability (UvPD).

It could then be deduced that in both scenarios we aim to maximize a utility reflected by our browses (a good internet service in the first case and suggestion of a relating product in the second) while keeping the privacy reflected by patterns shown in sequence of webpages visited and obscured by the underlying browses to a certain upper bound.

It could be argued that the best solution for such invasion of privacy problems for a user is to find approaches seeking tradeoff between some utility and the resulting privacy leakage.

One method to carry this out, is the use of website proxies. Such websites offer the user complete anonymity by coding the IP address of the visited web pages into a series of characters which can no longer be traced.

As an example if a user visits the proxy website www.proxy.com, they will see a URL box where they could enter the name of their intended web page. The proxy website will then open the new web page as an extension of its own URL, meaning in the address bar, there will be no trace of the desired web page. Furthermore, even if the eavesdropper attains the new link in the address bar, they cannot simply revisit the website as the new address is created only for the momentary accommodation of the user and will be destroyed as soon as the user is done with the URL.

The only detriment to such proxy websites is their limited bandwidth and time-outs. As a solution, we can employ multiple proxy websites and thus cut down on the utility loss. However, by using multiple proxy sites we are allowing a level of privacy breach into our browses where the eavesdropper is able to deduce some information about our tendencies by observing the distribution of used webpages over multiple networks. Then, if a utility function based upon connection speed -bandwidth- for the user is calculable, a privacy constrained problem between the user and an eavesdropper (for example a service provider) could be defined.

To formalize such trade-off problems, we propose a partitioning framework in which a set of M items (e.g. websites) are placed into N bins (e.g. proxy sites) to maximize certain payoff function while meeting constraints imposed due to concerns of privacy leakage or distinguishing capability. Then, a given distribution $G_1, G_2, ...$ on the random items represents a particular operating mode or behavior over a larger scale, which the user is willing to be classified. Under a particular distribution, however, a

user is more concerned of the sequence patterns, which could be inferred to certain extent by the resulting patterns of bins (e.g. proxy sites) deployed. On top of these two issues, a utility function for each bin is employed to reflect the throughput for this particular bin.

Our overarching goal is to pose and then solve the underlying constrained combinatorial optimization problems by first considering the tradeoff between meta leakage and binning utility, and then to further add the distinguishability into the objective by seeking optimal binning of items.

Due to the exponentially growing cost in searching such partitions, we instead seek sufficient conditions under which polynomial order complexity algorithms exist by resorting to multi-submodular and submodular structures in our problems.

A. Related Works

The concept of privacy has already been explored in many works such as [1], [2] where a general but non-mathematical explanation was offered. However; in our work, we go into further details as to what privacy represents in our framework and how it could be formulated into many settings.

As for the addition of distinguishability, to the best of our knowledge, the closest work to ours was done by [3] where they considered the tradeoff between distinguishability and information leakage when the former is quantified using KL-distance and the latter using mutual information. Lately, [4] also considered a problem close to ours where they used the concept of missed detection and false alarm as the basis of hypotheses testing. However, in both works, the authors were interested in a randomized binning process which tended to be more complex than our routine. In both cases, this came at the cost of sacrificing the concept of trade-off by only allowing very small privacy leakage (less than ϵ) in order to invoke a first order approximation to reduce the computational complexity of the optimization problems. However; in our work, we are concerned with any amount of trade-off with deterministic binning to seek suboptimal and polynomial order approximations enabled by multi-submodular set function conditions.

The multi-submodular function was discussed extensively in [5] where they introduced a series of sufficient conditions on multi-submodular set functions by which the multi-submodular problem could be transformed into a submodular set function problem. Then, further discussions about the existence of a solution to the new problem were made. By doing so -and if a solution were proven to exist-, the complexity of the problem could be shown to be reduced from NP to polynomial. Such sufficient conditions have been adopted by us in seeking multi-submodular solutions in our problem settings.

This paper represents an expansion of our previous works in [6], [7] where in [6] we had a goal of finding the average time required to detect one of multiple non-overlapping specific subgraphs. This goal was carried out by finding the output of the network to a set of pre-specified queries.

In this paper, we shift our attention from detection of an active subgraph to seeking tradeoff between optimizing utility of partitioning a set of random items and restraining information leakage.

B. Paper Contributions

It is important to note that some of the problems introduced in this paper, were partially considered in our previous work [8]. However, due to page limitations, the mathematical derivation of the results was left out.

As for the novel problem addressed in this paper, our goal is to offer (1) a measure of distinguishability between the $K = 2$ hypothesis; (2) a measure of average leaked information given any of the K hypothesis is active; (3) a formulation of a multi-agent multi-variant optimization problem with a privacy leakage constraint; (4) the development of insight into the complexity of such an NP-hard problem and further sufficient conditions under which we can simplify it into a polynomial problem; (5) a description of the algorithm utilized to find the solution given the sufficient conditions followed by the proximity results and finally (6) numerical examples to further demonstrate the applicability of submodular solutions, as compared with the results using exhaustive search in a small scale setting.

C. Paper Organization

The rest of this paper is organized as follows. In Section II we formulate the problems in terms of privacy and utility functions. We dissect what the goal and the constraints are. In Section III, we inspect the overall utility functions of each problem and then find the sufficient conditions under which they are equipped with multi-submodular property. In Section IV an algorithm for each case is developed offering a polynomial complexity and an accuracy level of e or $\frac{1}{e}$ depending on whether the goal is to minimize or maximize the overall utility function respectively. Finally, in Section VII, we conclude the chapter.

II. System Model

First, we propose an abstract framework to formalize the goal of seeking partition of M items into N bins. More specifically, we aim to allocate each one of $1 \leq i \leq M$ possible items (queries) to one of N output bins. There could be at most N^M such partitions. It follows that any set allocation $A_l, 1 \leq l \leq N^M$ results in N sets $S_j^{(l)} \subseteq \{1, 2, ..., M\}, j = 1, 2, ..., N$. Each such set is defined as

$$S_j^{(l)} = \{i | \theta_{i,j}^{(l)} = 1\}$$
$$where \quad \theta_{ij}^{(l)} = \begin{cases} 1 & i \in S_j^{(l)} \\ 0 & i \notin S_j^{(l)} \end{cases} \tag{1}$$

We further assume $S_j^{(l)} \cap S_k^{(l)} = \emptyset, j \neq k$. Furthermore we have $\bigcup_{j=1}^{N} S_j^{(l)} = \{1, 2, ..., M\}$. Finally the size of each such set $S_j^{(l)}$ is defined as $L_j^{(l)}$.

A. Probabilistic Model

We assume at any time slot one and only one of the inputs is chosen with a certain probability. Thus, assuming hypotheses G_p is active and if we use variable $X \in \{1, 2, ..., M\}$ as a representation of set of items, we could have $P(X = i | G_p) = P(\gamma_{i_p} = 1) = \pi_{i,p}, 1 \leq i \leq M, , p \in \{1, 2\}$ as a representation of the probability of choosing item i from the set X under hypotheses G_p

where $\gamma_{i_p} \in \{0,1\}$. It further follows that $\sum_{i=1}^{M} \gamma_{i_p} = 1, p \in \{1,2\}$, stipulating that one and only one of M items is selected.

Next, we introduce an observable random variable $Y \in \{1, 2, \cdots, N\}$, denoting the index of the bin (the proxy site) employed to carry one of the $M > N$ items. It follows that the probability of each bin's appearance given a set allocation scheme such as A_l under hypotheses G_p will be equal to

$$P(Y = j|A_l, G_p)$$
$$= \sum_{i=1}^{M} P(Y = j|A_l, X = i, G_p) P(X = i|A_l, G_p)$$

$$= \sum_{i=1}^{M} P(Y = j|A_l, X = i) P(X = i|G_p) \qquad (2)$$

Furthermore, $P(Y = j|A_l, X = i) = \theta_{ij}^{(l)} \in \{0,,1\}$. It thus follows that

$$P(Y = j|A_l, G_p) = \sum_{i=1}^{M} \theta_{ij}^{(l)} P(X = i|G_p) \rightarrow$$
$$P(Y = j|A_l, G_p) = \sum_{i \in S_j^{(l)}} \pi_{i_p} = \alpha_j^{(l,p)} \qquad (3)$$

B. Revealed Information

By choosing to allocate M items to N bins where $N \leq M$, we have injected ambiguity and uncertainty into the output binning index sequence about the input item sequence over a successive n visits or channel uses. In other words, if we originally chose to transmit n of such items, our total set of possible sequences would be of form $\overrightarrow{\mathbf{X^n}} = [X_1 X_2 ... X_n]$ out of M^n possible outcomes. From an observer's perspective which can only have access to which one of N bins is deployed in each time slot, sequences in the form of $\overrightarrow{\mathbf{Y^n}} = [Y_1 Y_2 ... Y_n]$ has cardinality of at most $N^n < M^n$. Despite the amount of uncertainty added due to the many-to-one mapping between items and bins, the output sequence sill reveals certain amount of information regarding the patterns of sequences of M random items.

This observation could be further studied by indicating how our allocation system resembles a coding framework where we have an equivalent channel whose input variable is X and output Y, as show in Figure 1.

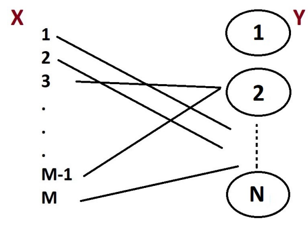

Figure 1. Coding Channel Representation of the Problem

Under such a framework, the equivalent channel output sequence $\overline{\mathbf{Y}}^n$ can help an eavesdropper classify the input sequence $\overline{\mathbf{X}}^n$ into a number of differential classes. As a result, information about the specific input item patterns is leaked to certain degree and can be measured using conditional mutual information $I(X;Y|A_l, G_p)$ between X and Y, under a hypotheses G_p given a particular binning (i.e. partition A_l relationship as illustrated in Figure 1.

Such conditional mutual information thus measures the maximum number of bits of meta information about item sequence per channel use. Therefore, we can have at most $2^{nI(X;Y|A_l,G_p)}$ sequences $\overline{\mathbf{X}^n}$ distinguishable by inferring based on $\overline{\mathbf{Y}^n}$. We thus adopt $I(X;Y|A_l, G_p)$ as the privacy metric conditioned on a particular partition mapping A_l under a specific hypotheses G_p. It follows that due to the combinatorial nature of a set allocation problem there are a total of N^M possible methods to allocate these M items to the N sets.

We can formulate the mutual information over a set allocation $A_l, 1 \leq l \leq N^M$ under hypotheses G_p as:

$$I(X;Y|A_l, G_p) = H(X|A_l, G_p) - H(X|Y, A_l, G_p) =$$
$$H(Y|A_l, G_p) - H(Y|X, A_l, G_p) = H(Y|A_l, G_p) =$$
$$H(\alpha_1^{(l,p)}, \alpha_2^{(l,p)}, ..., \alpha_N^{(l,p)}) \qquad (4)$$

where we have used the notion of $H(a_1, a_2, ..., a_m) = -\sum_{v=1}^{m} a_v \log a_v$ and the fact that $H(Y|X, A_l, G_p) = 0$ because if the input X and the channel scheme A_l and hypotheses G_p are known, then output Y will offer no uncertainty.

C. Distinguishability Measure

In this section, we aim to define a measure which can evaluate the distinguishability between a number of multinomial distributions.

Given that a certain set allocation A_l has been chosen, depending on the choice of G_p, two different distributions defined by using variables $\alpha_j^{(l,p)}$ could be developed where we have:

$$\alpha_j^{(l,p)} = P(Y = j|A_l, G_p), p \in \{1,2\}, j \in \{1, ..., N\} \quad (5)$$

where the two probability sets $P_1 = \{\alpha_j^{(l,1)}, j = 1, 2, ..., N\}$ and $P_2 = \{\alpha_j^{(l,2)}, j = 1, 2, ..., N\}$ represent the multinomial distributions under allocation scheme l given two prior probability distributions G_1 and G_2 which we aim to distinguish. To evaluate the distinguishability between two such distributions, we utilize the definition of symmetric KL divergence. We would have the measure defined as:

$$\frac{1}{2}\{KL(P_1||P_2) + KL(P_2||P_1)\}$$
$$= \frac{1}{2} \sum_{j=1}^{N} (\alpha_j^{(l,1)} - \alpha_j^{(l,2)})(\log \alpha_j^{(l,1)} - \log \alpha_j^{(l,2)}) \qquad (6)$$

We could now use the above definition in developing our problem formulation.

It is important to note how this measure of distinguishability is different in nature than that of leaked information. Through binning, a sequence of outputs are produced which we aim to limit. However, we could use an average (note the average nature of KL distance) of these sequences over multiple choices as a measure of maximum likelihood detection deciding which hypotheses is truly active. By

doing so, we are both controlling the level of leaked information and utilizing it for further gain.

D. Problem Formulation

Using the definitions addressed above, we could show that the UvP problem could be modeled in the form of finding the allocation scheme A_l which satisfies [8]

$$\boxed{\max_{0\leq l\leq N^M} \sum_{j=1}^{N} [\alpha_j^{(l,1)} f_j(S_j^{(l,1)}) + \lambda\alpha_j^{(l,1)} \log(\alpha_j^{(l,1)})]} \quad (7)$$

where $f_j(S_j^{(l,1)})$ represents the utility function offered by the j^{th} bin while using allocation scheme A_l under G_p hypotheses and $\lambda \geq 0$ demonstrates the Lagrange multiplier accounting for the leaked information constraint. Equation (7) is simply a sum of functions defined over a series of sets. We refer to these as multi-variate set functions seeing as how their values are based upon specific sets and variables introduced in each of these sets. Furthermore, we note that since there is only one active hypotheses G_1 in this scenario, we can have $\alpha_j^{(l,1)} = \alpha_j^{(l)}$ and $f_j(S_j^{(l,1)}) = f_j(S_j^{(l)})$.

We then aim to distinguish these two distributions. We thus choose binning as our method of observation meaning we attempt to find an M-to-N mapping between the set of M items and set of N outputs which can best help us distinguish which multinomial distribution was chosen.

Overall, we aim to find a partition of M items under which we can maximize the binning utility of the solution while keeping the distinguishability between G_1 and G_2 and the leaked information about the item probabilities higher and less than some thresholds respectively.

We next pose our problem as a constrained optimization problem:

$$\boxed{\begin{aligned} &\max_{1\leq l\leq N^M} \sum_{p=1}^{2}\sum_{j=1}^{N} P(G_p)[\alpha_j^{(l,p)} f(S_j^{(l,p)}) \\ &\quad +\lambda_2\alpha_j^{(l,p)} \log\alpha_j^{(l,p)}] + \lambda_2 I_{th} \\ &+\frac{\lambda_1}{2}\sum_{j=1}^{N} (\alpha_j^{(l,1)} - \alpha_j^{(l,2)})(\log\alpha_j^{(l,1)} - \log\alpha_j^{(l,2)}) \quad (8)\end{aligned}}$$

holds true. Furthermore, $\lambda_1, \lambda_2 > 0$ are Lagrange multipliers accounting for the constraints over distinguishability and leaked information respectively.

III. Multi-Submodular Set Functions as a Means of Solution

As mentioned previously, both UvP and UvPD are NP-complicated (they are solved when searches over N^M and 2^M possible set allocations are carried out respectively and the best allocation is chosen). Still, we could opt to utilize the definition of multi-submodular set functions so as to reduce the complexity to that of polynomial at the cost of accuracy.

A. Imposing Multi-submodularity

In [5], it was shown that if we can prove multi-submodularity for functions such as those formulated in Eq. (7), then they could be modeled as simpler problems (submodular set functions). We thus, aim to find the sufficient conditions for such occurrence. To do so, we first offer a review of multi-submodularity.

As in [5], if we define $\mathbb{M} = \{1, 2, ..., M\}$, then a multi-variate function $F : (2^{\mathbb{M}})^N \to \mathbb{R}^+$ is multi-submodular if for all pairs of tuples $(S_1, ..., S_N)$ and $(T_1, ...T_N) \in (2^{\mathbb{M}})^N$ we will have:

$$\begin{aligned} F(S_1, ..., S_N) + F(T_1, ..., T_N) &\geq F(S_1 \cup T_1, ..., S_N \cup T_N) \\ &+F(S_1 \cap T_1, ..., S_N \cap T_N) \quad (9)\end{aligned}$$

Since in our formulation functions are separately defined on different sets, the condition in Eq. (9) is simplified to the sufficient condition of submodularity of $F_j^{(l)} = \alpha_j^{(l)} f_j(S_j^{(l)}) + \lambda\alpha_j^{(l)} \log(\alpha_j^{(l)})$ for all sets S_j. We thus need to find the sufficient condition for submodularity of $F_j^{(l)}$ when defined over a set S_j.

As for the UvPD problem, we already are aware that all we needed to impose was submodularity over the overall utility function.

B. Imposing Separate Submodularities for UvP

For an easier mathematical representation of the following derivation we denote $F(S_j) = F_j^{(l)}$. Furthermore, we denote $f_j(S_j^{(l)}) = f(S_j^{(l)})$. Both these denotations allude to the fact that once a set allocation A_l is chosen, its index could be dropped.

In the next step, we opt to use diminishing return property as the means of making certain each of these functions are submodular. Following is a definition of diminishing returns for submodular functions, after which we derive the sufficient conditions for the case discussed in UvP).

Diminishing Property Return dictates that if we define $\mathbb{S}$ as the universal set, a set function $F : 2^{\mathbb{S}} \to \mathbb{R}^+$ is submodular if, for all $A, B \subseteq \mathbb{S}$ with $A \subseteq B$ and for each $x \in \mathbb{S} - B$ we have [9]:

$$F(A \cup \{x\}) - F(A) \geq F(B \cup \{x\}) - F(B) \quad (10)$$

Now we attempt to expand Eq. (10) for each $F(S_j)$. However, to properly do so, we first need to account for the behavior of this function.

We have defined $F(S_j) = \alpha_j f(S_j) + \lambda\alpha_j \log(\alpha_j)$ where it seems that the function has a singular relationship with set S_j. However; there is a secondary relationship the function shares with the set $S_j^{\mathbf{C}} = S - S_j$ where $S = \mathbb{S}$ represents the universal set. This relationship could be modeled as $F(S_j^{\mathbf{C}}) = (1 - \beta_j)f(S - S_j^{\mathbf{C}}) + \lambda(1 - \beta_j) \log(1 - \beta_j)$ where we have used the fact that $\beta_j = 1 - \alpha_j$ seeing as how we define

$$\beta_j^{(l)} = \sum_{i\in\{S-S_j^{(l)}\}} \pi_i \quad (11)$$

Thus, for any set S_j, we must find the sufficient conditions for the existence of diminishing property for

both functions $F_1(S_j) = \alpha_j f(S_j) + \lambda\alpha_j \log(\alpha_j)$ and $F_2(S_j) = (1-\alpha_j) f(S-S_j) + \lambda(1-\alpha_j)\log(1-\alpha_j)$. To do so, we will evaluate their sufficient conditions and then find their intersection as the final conditions (assuming they do not negate one another).

Note: For any further references, we first need to address a series of variable and function definitions which are going to play a vital role in the rest of this chapter:

Definitions for Problem 1

1. Any variable represented with a capital Letter represents a set.
2. Any variable represented with a small letter represents an element.
3. $A - B$ represents a set containing all elements of set A which do not appear in set B.
4. α_x represents the probability of item x and α_A represents the sum of probabilities of items mapped into a set A.
5. α_{BA} represents the difference in the sum of probabilities of items mapped into the sets B and A which could be further shown as $\alpha_{BA} = \alpha_B - \alpha_A$.
6. $g(C, D)$ represents the 1^{st} order difference of a set function $f(C)$ from $f(C-D)$ where $D \subseteq C$ which could be formulated as $g(C,D) = f(C) - f(C-D)$.
7. $q(C, C_1, D, D_1)$ represents the 2^{nd} order difference of a set function $f(C)$ where $C_1 \subseteq C$ and $D_1 \subseteq D$ which could be formulated as $q(C, C_1, D, D_1) = g(C, D) - g(C_1, D_1)$.
8. We assume the probability of items is sorted in a decreasing manner such as $\pi_1 \geq \pi_2 \geq \ldots \geq \pi_M$.

Theorem III.1. *The set functions $F_1(S_j)$ and $F_2(S_j)$ and as a result $F(S_j)$ are submodular if*

(1) $g(S_j, S_w) \leq 0$

(2) $q(S_j, S_u, S_w, S_y) \leq 0$

(3) $|g(S_j, S_w)| \geq \lambda \log(\frac{1}{\pi_M})$

for all possible sets $S_w \subseteq S_j \subseteq S$ and $S_u \subseteq S_j$ and $S_y \subseteq S_w$ where S is the universal set.

The proof for this theorem is presented in the Appendix under Theorem III.1. In the proof, a series of sufficient conditions for either $F_1(S_j)$ and $F_2(S_j)$ are evaluated separately. This is done because although their conditions turn out to be the same, their derivations are vastly different and require separate discussions. We then acquire the more restrictive set of sufficient conditions between the two sets of sufficient conditions for either $F_1(S_j)$ or $F_2(S_j)$. In the proof, we rewrite inequality (10) for set function $F_1(S_j)$, start factorizing α_A, α_{BA} separately and α_x and 1 together and impose sufficient conditions so that each of their coefficients is always positive.

Unfortunately, [5] does not turn a multi-submodular problem to a submodular one; they simply prove that this could be done. Thus, in order to expand upon the idea of polynomial complexity of solution algorithms we opt to assume $N = 2$ and offer the reader the algorithm to deal with such a specific case. We then calculate the complexity imposed by the algorithm to further stress the benefits of using such an idea at the expense of accuracy.

C. Specific Case of $N = 2$

As mentioned previously, in order to show the applicability of submodular functions we choose to reiterate the overall utility function dictated in UvP for when $N = 2$:

$$\begin{aligned} \max_{1\leq l \leq 2^M} \quad & \alpha_1^{(l)} f(S_1^{(l)}) + (1-\alpha_1^{(l)}) f(S - S_1^{(l)}) \\ &+\lambda(I_{th} + \alpha_1^{(l)} \log(\alpha_1^{(l)}) + (1-\alpha_1^{(l)})\log(1-\alpha_1^{(l)})) \\ &= T(S_1) \end{aligned} \tag{12}$$

As can be seen, the problem is still exponentially complex seeing as how we need to search over 2^M possible solutions to find the optimal. Thus, once again we aim to impose multi-submodularity (in this case simplified to submodularity) on the new utility function. For the utility function above the same results derived for a general N could be used as a set of sufficient conditions. However, taking into account the joint relationship between the 2 sets and writing the same Inequality (10) for Equality (12) we are able to find a less restrictive set of sufficient conditions for the submodularity of this utility function as indicated below:

Theorem III.2. *When $N = 2$, the function in Eq. (12) is submodular if*

(1) $g(S_j^{(l)}, S_w^{(l)}) \leq 0$

(2) $2|g(S_j^{(l)}, S_w^{(l)})| \geq \lambda \log(1 + \frac{1-\pi_M}{\pi_M^2})$

(3) $q(S_j^{(l)}, S_u^{(l)}, S_w^{(l)}, S_y^{(l)}) \leq 0$

for all possible sets $S_w^{(l)} \subseteq S_j^{(l)} \subseteq S$ and $S_u^{(l)} \subseteq S_j^{(l)}$ and $S_y^{(l)} \subseteq S_w^{(l)}$ where S is the universal set.

where we have deduce that the new set of sufficient conditions are less restrictive because they allow a smaller lower bound (a looser lower bound) for the absolute value of the first order derivative of the set function $f(S_j)$.

The proof for this theorem is presented in Appendix under theorem III.2. In the proof, we rewrite inequality (10) for set function described in Eq. (12), start factorizing α_A, α_{BA} separately and α_x and 1 together and impose sufficient conditions so that each of their coefficients is always positive.

D. Imposing Separate Submodularities for the UvPD Problem

In this section, we aim to find sufficient conditions for multisubmodularity of the overall utility function described for the UvPD Problem.

Once again, we use the intuition offered in the previous section meaning we aim to break down Eq.(8) into a sum of simpler set functions and then impose submodularity over all such sets and thus guarantee the overall multisubmodularity as well. Following the previous reasoning, we aim to find the sufficient conditions of submodularity for

$$\begin{aligned} F(S_j^{(l)}) &= P(G_1)[\alpha_j^{(l,1)} f(S_j^{(l,1)}) + \lambda_2 \alpha_j^{(l,1)} \log \alpha_j^{(l,1)}] \\ &+P(G_2)[\alpha_j^{(l,2)} f(S_j^{(l,2)}) + \lambda_2 \alpha_j^{(l,2)} \log \alpha_j^{(l,2)}] \\ &+\frac{\lambda_1}{2}(\alpha_j^{(l,1)} - \alpha_j^{(l,2)})(\log \alpha_j^{(l,1)} - \log \alpha_j^{(l,2)}) \end{aligned} \tag{13}$$

for all sets S_j. Assuming that a specific allocation scheme A_l has been chosen, we can further simplify Eq.(14) by rewriting

$$F(S_j^{(l)}) = F(S_j) = P(G_1)[\alpha_j^{(1)} f(S_j^{(1)}) + \lambda_2 \alpha_j^{(1)} \log \alpha_j^{(1)}] \\ +P(G_2)[\alpha_j^{(2)} f(S_j^{(2)}) + \lambda_2 \alpha_j^{(2)} \log \alpha_j^{(2)}] \\ +\frac{\lambda_1}{2}(\alpha_j^{(1)} - \alpha_j^{(2)})(\log \alpha_j^{(l,1)} - \log \alpha_j^{(l,2)}) \tag{14}$$

where we have simply dropped every iteration of l from our previous formulation for easier representation.

As was the case previously, it seems that the function $F(S_j)$ has a singular relationship with set S_j. However; there is a secondary relationship the function shares with the set $S_j^{\mathbf{C}} = S - S_j$ where $S = \mathbb{S}$ represents the universal set. This relationship could be modeled as

$$F(S_j^C) = P(G_1)[(1 - \beta_j^{(l,1)}) f(S - S_j^{(l,1)^C}) \\ +\lambda_2(1 - \beta_j^{(l,1)}) \log (1 - \beta_j^{(l,1)})] \\ +P(G_2)[(1 - \beta_j^{(l,2)}) f(S - S_j^{(l,2)^C}) \\ +\lambda_2(1 - \beta_j^{(l,2)}) \log (1 - \beta_j^{(l,2)})] \\ +\frac{\lambda_1}{2}(\beta_j^{(l,2)} - \beta_j^{(1)})(\log (1 - \beta_j^{(1)}) - \log (1 - \beta_j^{(2)})) \tag{15}$$

where we have used the fact that $\beta_j = 1 - \alpha_j$ seeing as how we define

$$\beta_j^{(l)} = \sum_{i \in \{S - S_j^{(l)}\}} \pi_i \tag{16}$$

Thus, for any set S_j, we must find the sufficient conditions for the existence of diminishing property for both functions described in Eq.(14) and Eq.(15). To do so, we will evaluate their sufficient conditions and then find their intersection as the final conditions (assuming they do not negate one another).

Theorem III.3. *The set functions defined for the UvPD problem is submodular if*

$$(1) \quad q(S_j, S_w, S_r, S_t) \leq 0 \\ (2) \quad g(S_j, S_w) \leq 0 \\ (3) \quad |g(S_j, S_w)| \geq \\ max\{\lambda_2 \log \omega_1' + \frac{\lambda_1}{P(G_1)} \log (\omega_1' \omega_2'), \\ \lambda_2 \log \omega_2' + \frac{\lambda_1}{P(G_2)} \log (\omega_2' \omega_1')\} \\ , \omega_1' = \frac{1}{\pi_{M_1}}, \omega_2' = \frac{1}{\pi_{M_2}}$$

for all possible sets $S_w \subseteq S_j$ and $S_t \subseteq S_r$ and $S_r \subseteq S_j$.

The proof for this theorem is presented in the appendix under Theorem III.3. Once again in the proof, a series of sufficient conditions for $F(S_j)$ as formulated in either Eq.(14) or in Eq.(15) are evaluated separately. This is done because although their conditions turn out to be the same, their derivations are vastly different and require separate discussions. It follows that since both functions require the same set of sufficient conditions, the overall function $F(S_j)$ which represents either representation being chosen, also follows the same set of sufficient conditions.

It is important to note how adding distinguishability as a factor to account for has changed the dynamic of our sets of sufficient conditions to assure accessibility of a submodular solution for the problem. In the UvP formulation of the problem, λ referred to the Lagrange multiplier accounting for the leaked information constraint. The same role is played by λ_2 in the formulation of UvPD problem. Furthermore, in both problems, we aim to maximize a utility while keeping the leaked information to a threshold. This would mean that if there were no concerns of distinguishability in the UvPD problem, we would be expecting the same sets of sufficient condition as the two problems basically become the same. As can be seen, if we assume λ_1 (which represents the Lagrange multiplier accounting for the distinguishability measure) is very small, then the two sets of sufficient conditions turn out to be the same (noting that there is basically only one hypotheses to account for when distinguishability is of no importance).

IV. Submodular Solution

In this section, we offer two algorithms specifying how each UvP and UvPD could be solved in a polynomial manner (when $N = 2$). We further expand upon how the complexity solutions has been compromised. Finally, we offer an example to showcase the relationship between solution error and complexity.

A. Solution to UvP

It is important to note that while all these papers dealt with the issue of constraints, they assumed much more complex constraints than we are dealing with in this discussion. Our only constraint is that $\alpha_1^{(l)} \leq \alpha_s$ where we assume $h(\alpha_s) = I_{th}, \alpha_s \leq 0.5$ which is obviously a down-monotone constraint. Thus, we can simply use the results from their work to create our own algorithm to find the submodular function solution to our problem. We present:

Algorithm 1: Submodular Function Solution to the UvP Problem)

1. Let $S_1 = argmax_{e \in X = \{1,\dots,M\}} T[S_1 = \{e\}]$ while $|0.5 - \alpha_1^{(l,1)}| \geq |0.5 - \alpha_s|$.
2. If there is an element $e \in X \setminus S_1$ such that $T[S_1 + \{e\}] \geq T[S_1]$ and $|0.5 - \alpha_1^{(l)} - \alpha_e| \geq |0.5 - \alpha_s|$, let $S_1 = S_1 + \{e\}$.
3. If there is an element $e \in S_1$ such that $T[S_1 \setminus \{e\}] \geq T[S_1]$ and $|0.5 - \alpha_1^{(l)} + \alpha_e| \geq |0.5 - \alpha_s|$, let $S_1 = S_1 - \{e\}$. Go to Step 2.
4. Return maximum of $T[S_1]$ and $T[X \setminus S_1]$.

B. Solution to the UvPD Problem

In this case our only constraint is that $\overline{\alpha_l} \leq \alpha_s$ where we assume $h(\alpha_s) = I_{th}, \alpha_s \leq 0.5$. Following the same pattern as that of Algorithm 1, we can write:

Algorithm 2: Submodular Function Solution to the UvPD Problem

1. Let $S_1 = argmax_{e \in X = \{1,\dots,M\}} T[S_1 = \{e\}]$ while

$U(S_1 = \{e\})$ and $W(S_1 = \{e\})$ both satisfy the constraints.

2. If there is an element $e \in X \setminus S_1$ such that $T[S_1 + \{e\}] \geq T[S_1]$ and $U(S_1 + \{e\})$ and $W(S_1 + \{e\})$ both satisfy the constraints, let $S_1 = S_1 + \{e_1\}$.
3. If there is an element $e \in S_1$ such that $T[S_1 \setminus \{e\}] \geq T[S_1]$ and $U(S_1 - \{e\})$ and $W(S_1 - \{e\})$ both satisfy the constraints, let $S_1 = S_1 - \{e\}$. Go to Step 2.
4. Return maximum of $T[S_1]$ and $T[X \setminus S_1]$.

In both cases, we know that at the very last step $T[S_1] = T[X \setminus S_1]$. Now we opt to calculate the complexities of this method. Steps 2 and 3 could repeat $(M-1)+(M-2)+ ... +1 = \frac{M(M+1)}{2}$ times each while every item could be removed and thus replaced a total of $2M$ times. Thus the total complexity of steps 2 and 3 is equal to $M^2(M+1) = O(M^3)$. The complexity of step 1 is also equal to M. Thus the total complexity of the solution is equal to $O(M^3)$.

These polynomial solutions simply make certain the maximal and minimal functions obtained are at least 0.432 and at most 2.315 times the optimal objective functions respectively. This range of error occurs because in this method, we are removing and adding members from and to the set S_1 one by one. Thus, at each decision point we are making one locally optimal decision. However, it is widely known that a locally greedy method is not necessarily globally optimal [10].

Finally, we find it important to note that in both these algorithms we are considering the case when $N = 2$. This is because while the studies for general multi-agent multi-variable solutions are still ongoing, there has not yet been any closed form solution algorithm for a general value of N. This is due to combinatorial nature of such binning problems; when $N = 2$, an item could either belong to one bin or the other and thus the idea of adding an item to a bin or removing it from a bin (as portrayed in either of the algorithms) is justifiable. However, when $N > 2$, there might be cases when no new item is added to a bin and instead any other two of the bins gain or lose items resulting in confusion while embedding an algorithm.

V. Examples for the UvP and UvPD Problem

In this section we aim to offer the reader an example where we aim to use the results gathered in previous theorems and Algorithms to first find a proper utility function given the specifics and then run and test the Algorithms' results with that of an exhaustive search to compare the two methods in terms of complexity and exactness of the solution. As a general assumption, we assume $M = 6, N = 2, \lambda = \lambda_2 = 0.4, \lambda_1 = 0.6, I_{th} = 0.47$ and that $f(S) = f(|S|)$ where S and $|S|$ represent any set and its cardinality respectively. We further assume the two following item probability distributions:

$$\begin{aligned} &\pi_{1,1} = 0.3, \pi_{2,1} = 0.25, \pi_{3,1} = 0.25, \\ &\pi_{4,1} = 0.10, \pi_{5,1} = 0.05, \pi_{6,1} = 0.05, \\ &\pi_{1,2} = 0.2, \pi_{2,2} = 0.2, \pi_{3,2} = 0.2, \\ &\pi_{4,2} = 0.15, \pi_{5,2} = 0.15, \pi_{6,2} = 0.10 \end{aligned} \tag{17}$$

It is important to note why we have chosen a rather small amount of the value of M; we have done so to be able to carry out the exhaustive search for the optimal set selection by hand without much difficulty. Thus, the choice of a small value of M in no manner indicates a limitation of the overall results. One of the simplest utility functions which could satisfy the conditions presented in theorem 4.3.2 is a quadratic function in the form of $f_p(|S|) = a|S|^2 + b|S| + C, p = 1, 2$. By calculating the 1^{st} and 2^{nd} order derivatives we can see that they follow the form of $g_p(|S|) = 2a|S| + b - a$ and $q_p(|S|) = 2a$ respectively. We could then show that the function $f_p(|S|) = -|S|^2 + |S| + 44, p = 1, 2$ satisfies all the sets of sufficient condition for imposing submodularity on both hypotheses.

Note: In our current and following derivations of a desirable utility function we focus on developing a quadratic utility function. Such an assumption might appear to be extremely limiting to our class of functions at first sight. However, as is further explained in [11], such an assumption is quite understandable and rather desirable seeing as how in many economic models, functions are written as extensions of quadratic function and thus later calculations are simplified while still maintaining the essence of a utility function.

A. Solution Comparison

The results of running Algorithm 1 on each hypotheses have been gathered in Table I. Each cell represents the maximum overall utility achieved in either hypotheses by each method where it is obvious that the probability distribution plays a more vital role on the exactness of the solution compared to the utility function -which is the same in both cases. Also of interest is the loss of utility at a desirable complexity reduction from NP to P which helps justify our persistence on utilizing submodular solutions.

Table I
Solution Exactness Comparison for UvP

Hypotheses	exhaustive search solution	submodular solution
1	32.8124	25.8124
2	25.8124	25.8124

Although we previously mentioned the possibility of a loss of accuracy, we still need to discuss the reason behind this disparity of results. While running the submodular algorithm for the first hypotheses, we need to both satisfy the leaked information constraint and maximize the utility function, this then results in the selection of the item with highest allowed probability ($\pi_{4,1} = 0.1$). However, once this item is fixed into the solution set, we cannot add any more items, because the information leakage threshold is reached and we cannot remove any item because then the utility is lessened. However, if we were to run the exhaustive search method, we would see that choosing the two items with the lowest probabilities ($\pi_{5,1} = \pi_{6,1} = 0.05$) would have offered a better overall utility. Thus, the biggest limitation in submodular set solutions appears in the process of allocating the first item to the solution set.

In the second hypotheses, such a problem does not occur since the first set solution choice already represents the best first choice.

VI. Examples for the UvPD Problem

Using the same item probability distributions as those in Eq.(17), we can see that the major difference is in how this time the average leaked information needs to be less than the desired threshold. As a result, we can showcase two cases of the scenario with two random hypotheses where the submodular solution varies in accuracy; if we assume $P(G_1) = 0.99, P(G_2) = 0.01$, we are basically offering a larger weight to the problematic scenario which results in a compromise in accuracy. Similarly, if we assume $P(G_2) = 0.99, P(G_1) = 0.01$, we are allowing more weight to the hypotheses with higher accuracy thus resulting in an optimal solution. We could thus deduce that accuracy in this class of problems has to do with both the hypotheses probabilities and each hypotheses' item probability distributions. We follow through with two such hypotheses which result in one uniform utility function $f_p(|S|) = -|S|^2 - 458|S| + 2796$. The results of these two hypotheses probabilities are gathered in the following table:

A. Utility function for Table II

Table II
SOLUTION EXACTNESS COMPARISON FOR UVPD

Hypotheses Selection	exhaustive search solution	submodular solution
$P(G_1) = 0.99$	1042.1	757.3537
$P(G_1) = 0.01$	757.3537	757.3537

It is important to note that other than each hypotheses' probability another factor which could play a vital solution in the accuracy of the results is the value λ_1 which intensifies the distinguishability measure. In our examples, we specifically chose a rather small value for λ_1 so as to be able to witness the role of each hypotheses. However, if λ_1 was a non-negligible value, such a straightforward could not be deduced.

VII. Conclusion and Future Works

In this paper, we introduced and formulated two problems widely regarded in online browses. While the first problem was briefly introduced in our previous work [8], we used this paper to expand upon the derivation of previous results. Furthermore, we expanded upon our previous problem by introducing the concept of hypotheses distinguishability thus realizing a fully novel problem. To carry out such goals, we formulated two separate utility function based upon the user's utilization of the network. We showcased how each problem formulation results in a multi-agent multi-variate problem which is NP-complicated. We then introduced the concept of submodularity and multi-submodularity which help reduce the complexity of such problems to that of a polynomial at the cost of some accuracy. We derived a series of sufficient conditions which would guarantee the existence of a solution. Once the existence of such solutions was guaranteed, we introduced algorithms that could help us reduce the complexities while refraining an acceptable solution. Finally, we used examples to illustrate trade-off between accuracy and complexity of the solutions.

References

[1] F. Laforet, E. Buchmann, and K. B?hm, "Individual privacy constraints on time-series data," *Information Systems*, vol. 54, pp. 74 – 91, 2015.

[2] J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle, "Privacy in the internet of things: threats and challenges," *Security and Communication Networks*, vol. 7, no. 12, pp. 2728–2742, 2014.

[3] J. Liao, L. Sankar, V. Y. F. Tan, and F. du Pin Calmon, "Hypothesis testing in the high privacy limit," in *Allerton*. IEEE, 2016, pp. 649–656.

[4] ——, "Hypothesis testing under mutual information privacy constraints in the high privacy regime," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 4, pp. 1058–1071, April 2018.

[5] R. Santiago and F. B. Shepherd, "Multi-agent and multivariate submodular optimization," *CoRR*, vol. abs/1612.05222, 2016.

[6] F. Bayat and S. Wei, "Non-adaptive sequential detection of active edge-wise disjoint subgraphs under privacy constraints," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 7, pp. 1615–1625, July 2018.

[7] ——, "Sequential detection of disjoint subgraphs over boolean mac channels: A probabilistic approach," in *2016 IEEE Globecom Workshops (GC Wkshps)*, Dec 2016, pp. 1–6.

[8] ——, "Partition of random items: Tradeoff between binning utility and meta information leakage," in *2018 25th International Conference on Telecommunications (ICT)*, June 2018, pp. 356–360.

[9] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver, *Combinatorial Optimization*. New York, NY, USA: John Wiley & Sons, Inc., 1998.

[10] M. Resende and C. Ribeiro, "Greedy randomized adaptive search procedures," in *Handbook of Metaheuristics*, F. Glover and G. Kochenberger, Eds. Kluwer Academic Publishers, 2003, pp. 219–249.

[11] D. Johnstone and D. Lindley, "Meanvariance and expected utility: The borch paradox," *Statist. Sci.*, vol. 28, no. 2, pp. 223–237, 05 2013.

VIII. Appendices

A. Theorem III.1

Proof. For an easier understanding, our proof of the theorem is broken into two sections:

Sufficient Conditions for $\mathbf{F_1(S_j)}$:

Starting for $F_1(S_j)$, by rewriting Eq. (10) we will have:

$$\begin{aligned}\alpha_{BA}[-g(B+\{x\},B)+\lambda\log[\frac{1}{1+\frac{\alpha_x}{\alpha_A+\alpha_{BA}}}]]\\ +\alpha_x[-g(B+\{x\},A+\{x\})+\lambda\log[\frac{1}{1+\frac{\alpha_{BA}}{\alpha_A+\alpha_x}}]]\\ +\alpha_A[-q(B+\{x\},B,A+\{x\},A)+\lambda\log\frac{1+\frac{\alpha_{BA}}{\alpha_A}}{1+\frac{\alpha_{BA}}{\alpha_A+\alpha_x}}\geq 0\end{aligned} \tag{18}$$

where we have used the definition of functions f and g.

Since the above inequality needs to hold true for all possible sets of $A \subseteq B$, $x \notin B$, we aim to determine the maximal amount enforced by the above set of inequalities. The first factorization results in two inequalities: (1) $g(C, D) \leq 0$ and (2) $|g(C, D)| \geq \lambda max(\log(1+\frac{\alpha_x}{\alpha_B}))$ for all sets $D \subseteq C$. To find the maximum of such a limit, we need to impose the one item with highest probability to $\{x\}$ and assume the one lowest probability item to set B. Then the above inequality is maximized.

The second factorization results in two inequalities: (1) $g(C, D) \leq 0$ and (2) $|g(C, D)| \geq \lambda max(\log(1+\frac{\alpha_{BA}}{\alpha_A+\alpha_x}))$ for all sets $D \subseteq C$. To find the maximum of such a limit, we need to impose the item with

lowest probability to $\{x\}$ and that $\alpha_A = 0$ and then have $\alpha_{BA} = 1 - \alpha_x = \alpha_B$.

The third factorization could be simplified. The logarithm argument consists of a nominator greater than denominator thus resulting in the overall logarithm argument to be positive. We thus only need to impose that $q(C, C_1, D, D_1) \le 0$ for all sets $C_1 \subseteq C, D_1 \subseteq D$ and $D \subseteq C$.

It then follows that if the probability of each item is sorted in a decreasing manner as $\pi_1, ..., \pi_M$, we can write the set of 3 sufficient conditions for submodularity of set function $F_1(S_j), j = 1, ..., N$ as (1) $g(C, D) \le 0$ (2) $q(C, C_1, D, D_1) \le 0$ (3) $|g(C, D)| \ge \lambda \log [max(1 + \frac{\pi_1}{\pi_M}, 1 + \frac{1-\pi_M}{\pi_M})] = \lambda \log (\frac{1}{\pi_M})$ for all sets $C_1 \subseteq C, D_1 \subseteq D$ and $D \subseteq C$.

Sufficient Conditions for $\mathbf{F_2(S_j)}$: We now follow the same method for $F_2(S_j), j = 1, ..., N$ by rewriting Eq. (10):

$$\begin{aligned}
&\alpha_{BA}[-g(S-B, S-B-\{x\}) \\
&-\lambda \log (1 + \frac{\alpha_x}{1-\alpha_A-\alpha_{BA}-\alpha_x})] \\
&+\alpha_x[-g(S-A-\{x\}, S-B-\{x\}) \\
&-\lambda \log (1 + \frac{\alpha_{BA}}{1-\alpha_A-\alpha_x-\alpha_{BA}})] \\
&+\alpha_A[q(S-A, S-A-\{x\}, S-B, S-B-\{x\}) \\
&+\lambda \log (\frac{1-\alpha_A}{1-\alpha_A-\alpha_x}\frac{1-\alpha_A-\alpha_{BA}-\alpha_x}{1-\alpha_A-\alpha_{BA}})] \\
&+1[-q(S-A, S-A-\{x\}, S-B, S-B-\{x\}) \\
&+\lambda \log (\frac{1-\alpha_A-\alpha_x}{1-\alpha_A}\frac{1-\alpha_A-\alpha_{BA}}{1-\alpha_A-\alpha_{BA}-\alpha_x})] \ge 0 \quad (19)
\end{aligned}$$

where we have used the definition of functions f and g.

Since the above inequality needs to hold true for all possible sets of $A \subseteq B$, $x \notin B$, we aim to determine the maximal amount enforced by the above set of inequalities. The first factorization results in two inequalities: (1) $g(C, D) \le 0$ and (2) $|g(C, D)| \ge \lambda \log (1 + \frac{\alpha_x}{1-\alpha_B-\alpha_x})$ for all sets $D \subseteq C$. To find the maximum of such a limit, we need to impose the one item with highest probability to $\{x\}$ and assume that α_B is maximal while still less than $1 - \alpha_x$. This limit is imposed so that the denominator is not equal to 0. Then the lower bound for $g(C, D)$ will be equal to $\lambda \log (1 + \frac{\pi_1}{\pi_M})$.

The second factorization results in two inequalities: (1) $g(C, D) \le 0$ and (2) $|g(C, D)| \ge \lambda max(\log (1 + \frac{\alpha_{BA}}{1-\alpha_B-\alpha_x}))$ for all sets $D \subseteq C$. To find the maximum of such a limit, we once again need to impose the one item with highest probability to $\{x\}$ and assume that $\alpha_A = 0$ and α_B is maximal while still less than $1-\alpha_x$. This limit is once again imposed so that the denominator is not equal to 0. The secondary lower bound for $g(C, D)$ will be equal to $\lambda \log (1 + \frac{1-\pi_M-\pi_{M-1}}{\pi_M})$.

The third and forth factorization could be simplified. We could write them as

$$\begin{aligned}
&[-1 + \alpha_A]\{q(S-A, S-A-\{x\}, S-B, S-B-\{x\}) \\
&+\lambda \log \frac{1-\alpha_A-\alpha_x}{1-\alpha_A}\frac{1-\alpha_A-\alpha_{BA}}{1-\alpha_A-\alpha_{BA}-\alpha_x}\}+ \\
&+\alpha_A \lambda \log (1) \ge 0 \quad (20)
\end{aligned}$$

Furthermore, the logarithmic argument is always positive seeing as how the nominator of the fraction inside it is greater than the denominator. It once more thus follows that we merely need to impose $q(C, C_1, D, D_1) \le 0$ for all sets $C_1 \subseteq C, D_1 \subseteq D$ and $D \subseteq C$ to help above inequality hold true.

Thus 3 sufficient conditions for submodularity of $F_2(S_j), j = 1, ..., N$ are developed: (1) $g(C, D) \le 0$ (2) $q(C, C_1, D, D_1) \le 0$ (3) $|g(C, D)| \ge \lambda \log (\frac{1}{\pi_M})$ for all sets $C_1 \subseteq C, D_1 \subseteq D$ and $D \subseteq C$.

Finally, the intersection of the two sets of sufficient conditions for submodularity of $F_1(S_j)$ and $F_2(S_j)$ gives us the sufficient conditions for submodularity of $F(S_j)$ which turns out to be the same as either of theirs.

□

B. *theorem III.2*

Proof. For the specific case $N = 2$:

$$\begin{aligned}
F(A) = \alpha_A f(A) + (1-\alpha_A) f(S-A) + \alpha_A \log (\alpha_A) \\
+(1-\alpha_A)\log(1-\alpha_A) \quad (21)
\end{aligned}$$

In such a case, diminishing return property requires that

$$\begin{aligned}
&\alpha_{BA}\{-g(B+\{x\}, B) - g(S-B, S-B-\{x\}) \\
&+\lambda \log (\frac{\alpha_A+\alpha_{BA}}{\alpha_A+\alpha_{BA}+\alpha_x}\frac{1-\alpha_A-\alpha_{BA}-\alpha_x}{1-\alpha_A-\alpha_{BA}})\}+ \\
&\alpha_x\{-g(B+\{x\}, A+\{x\}) - g(S-A-\{x\}, S-B-\{x\}) \\
&+\lambda \log (\frac{\alpha_A+\alpha_x}{\alpha_A+\alpha_x+\alpha_{BA}}\frac{1-\alpha_A-\alpha_x-\alpha_{BA}}{1-\alpha_A-\alpha_x})\} \\
&+\alpha_A\{-q(B+\{x\}, B, A+\{x\}, A) \\
&+q(S-A, S-A-\{x\}, S-B, S-B-\{x\}) + \lambda \\
&\log (\frac{\alpha_A+\alpha_x}{\alpha_A+\alpha_x+\alpha_{BA}}\frac{1-\alpha_A-\alpha_x-\alpha_{BA}}{1-\alpha_A-\alpha_x}\frac{\alpha_A+\alpha_{BA}}{\alpha_A} \\
&\frac{1-\alpha_A}{1-\alpha_A-\alpha_{BA}})\} \\
&+\{-q(S-A, S-A-\{x\}, S-B, S-B-\{x\}) \\
&+\lambda \log (\frac{1-\alpha_A-\alpha_x}{1-\alpha_A-\alpha_x-\alpha_{BA}}\frac{1-\alpha_A-\alpha_{BA}}{1-\alpha_A})\} \\
&\ge 0 \quad (22)
\end{aligned}$$

where we have used the definition of functions f and g.

We aim to find sufficient conditions so that the above inequality can hold true. The first factorization is the coefficients of $\alpha_{BA} \ge 0$. We thus need to ascertain that the coefficient is also positive. To do so, we impose that

$$\begin{aligned}
-2g(C, D) + \lambda \log (\frac{\alpha_A+\alpha_{BA}}{\alpha_A+\alpha_{BA}+\alpha_x}\frac{1-\alpha_A-\alpha_{BA}-\alpha_x}{1-\alpha_A-\alpha_{BA}})\} \\
\ge 0 \quad (23)
\end{aligned}$$

for all possible sets $D \subseteq C$. We thus need to impose two sufficient conditions:

$$\begin{aligned}
&(1) \quad g(C,D) \leq 0, \forall C \\
&(2) \quad 2|g(C,D)| \geq \\
&max(|\lambda \log(\frac{\alpha_A + \alpha_{BA}}{\alpha_A + \alpha_{BA} + \alpha_x} \frac{1-\alpha_A-\alpha_{BA}-\alpha_x}{1-\alpha_A-\alpha_{BA}})\}|) \\
&= \lambda \log K_1
\end{aligned} \tag{24}$$

The second factorization is the coefficients of $\alpha_x \geq 0$. We once again need to ascertain that the coefficient is also positive. To do so, we impose that

$$\begin{aligned}
&(1) \quad g(C,D) \leq 0, \forall C \\
&(2) \quad 2|g(C,D)| \geq \\
&max(|\lambda \log(\frac{\alpha_A + \alpha_x}{\alpha_A + \alpha_x + \alpha_{BA}} \frac{1-\alpha_A-\alpha_x-\alpha_{BA}}{1-\alpha_A-\alpha_x})|) \\
&= \lambda \log K_2
\end{aligned} \tag{25}$$

for all sets $D \subseteq C$. We will later determine what $max\{K_1, K_2\}$ entails.

The 3^{rd} and 4^{th} factorizations could be merged together in the following manner:

$$\begin{aligned}
&[-1+\alpha_A]\{-q(S-A, S-A-\{x\}, S-B, S-B-\{x\}) \\
&+\log(\frac{1-\alpha_a-\alpha_x}{1-\alpha_A-\alpha_{BA}-\alpha_x}\frac{1-\alpha_A-\alpha_{BA}}{1-\alpha_A})\} \\
&+\alpha_A\{-q(S-A, S-A-\{x\}, S-B, S-B-\{x\}) \\
&-q(B+\{x\}, B, A+\{x\}, A) \\
&+\log(\frac{\alpha_A+\alpha_x}{\alpha_A}\frac{\alpha_A+\alpha_{BA}}{\alpha_A+\alpha_{BA}+\alpha_x})\}
\end{aligned} \tag{26}$$

We can show that both λ coefficients are always positive. Thus, all we need to impose to guarantee the 3^{rd} and 4^{th} factorizations do not cause any ambiguities, is $q(C, C_1, D, D_1) \leq 0$ for all sets $C_1 \subseteq C, D_1 \subseteq D$ and $D \subseteq C$.

Finally, we need to calculate both values of K_1 and K_2 and determine the choice of maximum between them. It could be seen that

$$K_1 = max \log(1 + \frac{\alpha_x}{\alpha_B(1-\alpha_B-\alpha_x)}) \tag{27}$$

over all possible combinations of α_B and α_x where $\alpha_B + \alpha_x \leq 1$. Thus, in order to find a global maximum over the surface entailed by α_B and α_x, we first need to check if the Hessian of the function is always negative (whether the function is concave). However, as it turns out, this is not necessarily true for all values of α_B and α_x. We thus suffice to finding an upper bound for K_1 which is easier to calculate. This action is justifiable since we are looking for sufficient conditions and thus they do not need to be tailor made tight.

$$1 + \frac{\alpha_x}{\alpha_B(1-\alpha_B-\alpha_x)} \leq 1 + \frac{\alpha_x}{\pi_M \alpha_B} = L_1 \tag{28}$$

It could then be proven that the new function L_1 is always concave and thus a global maximum could be calculated. This maximum occurs when $\alpha_x = \pi_1$ and $\alpha_B = \pi_M$ and thus we would have:

$$K_1 = \log(1 + \frac{\pi_1}{\pi_M^2}) \tag{29}$$

As for K_2's value, it could be seen that

$$K_2 = max \log(1 + \frac{\alpha_B - \alpha_A}{(\alpha_A+\alpha_x)(1-\alpha_B-\alpha_x)}) \tag{30}$$

over all possible combinations of α_B and α_x and α_A where $\alpha_B + \alpha_x \leq 1$ and $\alpha_A \leq alpha_B$. We can use a variable $\alpha_z = \alpha_A + \alpha_x$ and using the definition $\alpha_{BA} = \alpha_B - \alpha_A$, we can rewrite the definition of K_2 as:

$$K_2 = max \log(1 + \frac{\alpha_{BA}}{\alpha_z(1-\alpha_z-\alpha_{BA})}) \tag{31}$$

over all possible combinations of α_z and α_{BA} where $\alpha_B + \alpha_x \leq 1$. We have thus managed to transform the definition of K_2 into a definition that is very similar to that of K_1's. We then use the same reasoning to acknowledge that

$$K_2 = \log(1 + \frac{1-\pi_M}{\pi_M^2}) \tag{32}$$

Finally, we will have:

$$max\{K_1, K_2\} = \log(1 + \frac{1-\pi_M}{\pi_M^2}) \tag{33}$$

Finally, the new set of sufficient conditions could be summed up as (1) $g(C,D) \leq 0$ (2) $q(C, C_1, D, D_1) \leq 0$ (3) $2|g(C,D)| \geq \lambda \log(1 + \frac{1-\pi_M}{\pi_M^2})$ for all sets $C_1 \subseteq C, D_1 \subseteq D$ and $D \subseteq C$.

C. theorem III.3

By first imposing the diminishing returns property to $F(S_j)$ as formulated in Eq.(14) we see it necessary to have:

$$\begin{aligned}
&\alpha_{x_1}\{-P(G_1)g(S_B^{(1)}+\{x\}^{(1)}, S_A^{(1)}+\{x\}^{(1)}) \\
&-\lambda_1(\log(\alpha_{B_2}+\alpha_{x_2}) - \log(\alpha_{A_2}+\alpha_{x_2})) \\
&-(P(G_1)\lambda_2+\lambda_1)(\log(\alpha_{B_1}+\alpha_{x_1}) - \log(\alpha_{A_1}+\alpha_{x_1}))\} \\
&+\alpha_{x_2}\{-P(G_2)g(S_B^{(2)}+\{x\}^{(2)}, S_A^{(2)}+\{x\}^{(2)}) \\
&-\lambda_1(\log(\alpha_{B_1}+\alpha_{x_1}) - \log(\alpha_{A_1}+\alpha_{x_1})) \\
&-(P(G_2)\lambda_2+\lambda_1)(\log(\alpha_{B_2}+\alpha_{x_2}) - \log(\alpha_{A_2}+\alpha_{x_2}))\} \\
\\
&+\alpha_{BA_1}\{-P(G_1)g(S_B^{(1)}+\{x\}^{(1)}, S_B^{(1)}) \\
&-\lambda_1(\log(\alpha_{B_2}+\alpha_{x_2}) - \log\alpha_{B_2}) \\
&-(P(G_1)\lambda_2+\lambda_1)(\log(\alpha_{B_1}+\alpha_{x_1}) - \log\alpha_{B_1})\} \\
&+\alpha_{BA_2}\{-P(G_2)g(S_B^{(2)}+\{x\}^{(2)}, S_B^{(2)}) \\
&-\lambda_1(\log(\alpha_{B_1}+\alpha_{x_1}) - \log\alpha_{B_1}) \\
&-(P(G_2)\lambda_2+\lambda_1)(\log(\alpha_{B_2}+\alpha_{x_2}) - \log\alpha_{B_2})\} \\
&+\alpha_{A_1}\{-P(G_1)q(S_B^{(1)}+\{x\}^{(1)}, S_B^{(1)}, S_A^{(1)}+\{x\}^{(1)}, S_A^{(1)}) \\
&-\lambda_1\log(\frac{\alpha_{A_2}}{\alpha_{A_2}+\alpha_{x_2}}\frac{\alpha_{B_2}+\alpha_{x_2}}{\alpha_{B_2}}) \\
&-(P(G_1)\lambda_2+\lambda_1)\log(\frac{\alpha_{A_1}}{\alpha_{A_1}+\alpha_{x_1}}\frac{\alpha_{B_1}+\alpha_{x_1}}{\alpha_{B_1}})\} \\
&+\alpha_{A_2}\{-P(G_2)q(S_B^{(2)}+\{x\}^{(2)}, S_B^{(2)}, S_A^{(2)}+\{x\}^{(2)}, S_A^{(2)}) \\
&-\lambda_1\log(\frac{\alpha_{A_1}}{\alpha_{A_1}+\alpha_{x_1}}\frac{\alpha_{B_1}+\alpha_{x_1}}{\alpha_{B_1}}) \\
&-(P(G_2)\lambda_2+\lambda_1)\log(\frac{\alpha_{A_2}}{\alpha_{A_2}+\alpha_{x_2}}\frac{\alpha_{B_2}+\alpha_{x_2}}{\alpha_{B_2}})\} \geq 0
\end{aligned} \tag{34}$$

Now we attempt to find the sufficient conditions for Inequality (34) to hold true.

If we assume $q(S_j, S_w, S_r, S_t) \leq 0, S_w \subseteq S_j, S_t \subseteq S_r, S_r \subseteq S_j$ we will satisfy the positivity of α_{A_1}, α_{A_2} coefficients.

It could further be seen that if we define

$$\omega_1 = max(\frac{\alpha_{B_1} + \alpha_{x_1}}{\alpha_{A_1} + \alpha_{x_1}})$$
$$\omega_2 = max(\frac{\alpha_{B_2} + \alpha_{x_2}}{\alpha_{A_2} + \alpha_{x_2}}) \tag{35}$$

and

$$\chi_1 = max(\frac{\alpha_{B_1} + \alpha_{x_1}}{\alpha_{B_1}})$$
$$\chi_2 = max(\frac{\alpha_{B_2} + \alpha_{x_2}}{\alpha_{B_2}}) \tag{36}$$

then as long as $g(S_j, S_w) \leq 0, S_w \subseteq S_j$ and

$$\begin{gathered}|g| \geq \\ max\{\lambda_2 \log \omega_1 + \frac{\lambda_1}{P(G_1)} \log(\omega_1\omega_2), \\ \lambda_2 \log \omega_2 + \frac{\lambda_1}{P(G_2)} \log(\omega_2\omega_1), \\ \lambda_2 \log \chi_1 + \frac{\lambda_1}{P(G_1)} \log(\chi_1\chi_2), \\ \lambda_2 \log \chi_2 + \frac{\lambda_1}{P(G_2)} \log(\chi_2\chi_1)\}\end{gathered} \tag{37}$$

we will have satisfied the diminishing returns property inequality. In the next step we attempt to calculate the values of $\omega_1, \omega_2, \chi_1$ and χ_2. We further note that $\omega_i, i \in \{1, 2\}$ and $\chi_j, j \in \{1, 2\}$ represent the same set functions under different hypotheses, we could thus find the maximum values of either ω_i and χ_j and then simply change the hypotheses index to derive the other. In the rest of this proof, we aim to find the maximal values of ω_1 and χ_1.

It could be shown that

$$\omega_1 = 1 + \frac{\alpha_{BA_1}}{\alpha_{A_1} + \alpha_{x_1}} \tag{38}$$

Thus, the maximum value of ω_1 occurs when α_{BA_1} and $\alpha_{A_1} + \alpha_{x_1}$ are maximum and minimum respectively. This could occur when $\alpha_{x_1} = \pi_{M_1}$, $\alpha_{A_1} = 0$ and $\alpha_{BA_1} = 1 - \pi_{M_1}$. Then we would have:

$$\omega_1 = \frac{1}{\pi_{M_1}}$$
$$\omega_2 = \frac{1}{\pi_{M_2}} \tag{39}$$

Similarly, it could be shown that

$$\chi_1 = 1 + \frac{\alpha_{x_1}}{\alpha_{B_1}} \leq 1 + \frac{\pi_{1_1}}{\alpha_{B_1}} \tag{40}$$

Thus, the maximum value of χ_1 occurs when α_{B_1} is minimum. This could occur when $\alpha_{B_1} = \pi_{M_1}$. Then we would have:

$$\omega_1 = \frac{\pi_{1_1} + \pi_{M_1}}{\pi_{M_1}}$$
$$\omega_2 = \frac{\pi_{1_2} + \pi_{M_2}}{\pi_{M_2}} \tag{41}$$

Furthermore, we could see that as long as $M \geq 2$, we will have $\omega_1 \geq \chi_1$ and $\omega_2 \geq \chi_2$. We can thus rewrite the final set of sufficient conditions in the following manner:

$$\begin{gathered}(1) \quad q(S_j, S_w, S_r, S_t) \leq 0 \\ (2) \quad g(S_j, S_w) \leq 0 \\ (3) \quad |g(S_j, S_w)| \geq \\ max\{\lambda_2 \log \omega_1 + \frac{\lambda_1}{P(G_1)} \log(\omega_1\omega_2), \\ \lambda_2 \log \omega_2 + \frac{\lambda_1}{P(G_2)} \log(\omega_2\omega_1)\} \\ , \omega_1 = \frac{1}{\pi_{M_1}}, \omega_2 = \frac{1}{\pi_{M_2}}\end{gathered}$$

Now we aim to carry out the same process for $F(S_j)$ as formulated in Eq.(15). By imposing the diminishing returns property, we see it necessary to have:

$$\begin{gathered}\alpha_{x_1}\{-P(G_1)g(S - S_A^{(1)} - \{x\}^{(1)}, S - S_B^{(1)} - \{x\}^{(1)}) \\ -\lambda_1(\log(1 - \alpha_{A_2} - \alpha_{x_2}) - \log(1 - \alpha_{B_2} - \alpha_{x_2})) \\ -(P(G_1)\lambda_2 + \lambda_1) \\ (\log(1 - \alpha_{A_1} - \alpha_{x_1}) - \log(1 - \alpha_{B_1} - \alpha_{x_1}))\} \\ +\alpha_{x_2}\{-P(G_2)g(S - S_A^{(2)} - \{x\}^{(2)}, S - S_B^{(2)} - \{x\}^{(2)}) \\ -\lambda_1(\log(1 - \alpha_{A_1} - \alpha_{x_1}) - \log(1 - \alpha_{B_1} - \alpha_{x_1})) \\ -(P(G_2)\lambda_2 + \lambda_1) \\ (\log(1 - \alpha_{A_2} - \alpha_{x_2}) - \log(1 - \alpha_{B_2} - \alpha_{x_2}))\} \\ \\ \alpha_{BA_1}\{-P(G_1)g(S - S_B^{(1)}, S - S_B^{(1)} - \{x\}^{(1)}) \\ -\lambda_1(\log(1 - \alpha_{B_2}) - \log(1 - \alpha_{B_2} - \alpha_{x_2})) \\ -(P(G_1)\lambda_2 + \lambda_1) \\ (\log(1 - \alpha_{B_1}) - \log(1 - \alpha_{B_1} - \alpha_{x_1}))\} \\ +\alpha_{BA_2}\{-P(G_2)g(S - S_B^{(2)}, S - S_B^{(2)} - \{x\}^{(2)}) \\ -\lambda_1(\log(1 - \alpha_{B_1}) - \log(1 - \alpha_{B_1} - \alpha_{x_1})) \\ -(P(G_2)\lambda_2 + \lambda_1) \\ (\log(1 - \alpha_{B_2}) - \log(1 - \alpha_{B_2} - \alpha_{x_2}))\} \\ \\ +(1 - \alpha_{A_1})\{-P(G_1)q(S - S_A^{(1)}, S - S_A^{(1)} - \{x\}^{(1)}, \\ S - S_B^{(1)}, S - S_B^{(1)} - \{x\}^{(1)}) \\ -\lambda_1 \log(\frac{1 - \alpha_{A_2}}{1 - \alpha_{A_2} - \alpha_{x_2}} \frac{1 - \alpha_{B_2} - \alpha_{x_2}}{1 - \alpha_{B_2}}) \\ -(P(G_1)\lambda_2 + \lambda_1) \log(\frac{1 - \alpha_{A_1}}{1 - \alpha_{A_1} - \alpha_{x_1}} \frac{1 - \alpha_{B_1} - \alpha_{x_1}}{1 - \alpha_{B_1}})\} \\ +(1 - \alpha_{A_2})\{-P(G_2)q(S - S_A^{(2)}, S - S_A^{(2)} - \{x\}^{(2)}, \\ S - S_B^{(2)}, S - S_B^{(2)} - \{x\}^{(2)}) \\ -\lambda_1 \log(\frac{1 - \alpha_{A_1}}{1 - \alpha_{A_1} - \alpha_{x_1}} \frac{1 - \alpha_{B_1} - \alpha_{x_1}}{1 - \alpha_{B_1}}) \\ -(P(G_2)\lambda_2 + \lambda_1) \log(\frac{1 - \alpha_{A_2}}{1 - \alpha_{A_2} - \alpha_{x_2}} \frac{1 - \alpha_{B_2} - \alpha_{x_2}}{1 - \alpha_{B_2}})\} \\ \geq 0\end{gathered} \tag{42}$$

Once more, it could further be seen that if we define

$$\omega_1' = max(\frac{1 - \alpha_{A_1} - \alpha_{x_1}}{1 - \alpha_{B_1} - \alpha_{x_1}})$$
$$\omega_2' = max(\frac{1 - \alpha_{A_2} - \alpha_{x_2}}{1 - \alpha_{B_2} - \alpha_{x_2}}) \tag{43}$$

and

$$\chi'_1 = max(\frac{1-\alpha_{B_1}}{1-\alpha_{B_1}-\alpha_{x_1}})$$
$$\chi'_2 = max\frac{1-\alpha_{B_2}}{1-\alpha_{B_2}-\alpha_{x_2}}) \tag{44}$$

then as long as $g(S_j, S_w) \leq 0, S_w \subseteq S_j$ and

$$\begin{aligned}
&|g| \geq \\
&max\{\lambda_2 \log \omega'_1 + \frac{\lambda_1}{P(G_1)} \log{(\omega'_1\omega'_2)}, \\
&\lambda_2 \log \omega'_2 + \frac{\lambda_1}{P(G_2)} \log{(\omega'_2\omega'_1)}, \\
&\lambda_2 \log \chi'_1 + \frac{\lambda_1}{P(G_1)} \log{(\chi'_1\chi'_2)}, \\
&\lambda_2 \log \chi'_2 + \frac{\lambda_1}{P(G_2)} \log{(\chi'_2\chi'_1)}\}
\end{aligned} \tag{45}$$

we will have satisfied the diminishing returns property inequality. In the next step we attempt to calculate the values of $\omega'_1, \omega'_2, \chi'_1$ and χ'_2. We further note that $\omega'_i, i \in \{1,2\}$ and $\chi'_j, j \in \{1,2\}$ represent the same set functions under different hypotheses, we could thus find the maximum values of either ω'_i and χ'_j and then simply change the hypotheses index to derive the other. In the rest of this proof, we aim to find the maximal values of ω'_1 and χ'_1.

It could be shown that

$$\chi'_1 = 1 + \frac{\alpha_{x_1}}{1-\alpha_{B_1}-\alpha_{x_1}} \tag{46}$$

Thus, the maximum value of ω_1 occurs when α_{x_1} and $\alpha_{B_1} + \alpha_{x_1}$ are both maximum. This could occur when $\alpha_{x_1} = \pi_{1_1}$, $\alpha_{B_1} + \alpha_{x_1} = 1 - \pi_{M_1}$. Then we would have:

$$\chi'_1 = \frac{\pi_{1_1} + \pi_{M_1}}{\pi_{M_1}}$$
$$\chi'_2 = \frac{\pi_{1_2} + \pi_{M_2}}{\pi_{M_2}} \tag{47}$$

Similarly, it could be shown that

$$\begin{aligned}
\omega'_1 = \frac{1-\alpha_{A_1}-\alpha_{x_1}}{1-\alpha_{B_1}-\alpha_{x_1}} &\leq \frac{1-\alpha_{x_1}}{1-\alpha_{B_1}-\alpha_{x_1}} \\
&\leq \frac{1-\pi_{M_1}}{1-\alpha_{B_1}-\alpha_{x_1}}
\end{aligned} \tag{48}$$

Thus, the maximum value of ω'_1 occurs when $\alpha_{B_1} + \alpha_{B_1} = 1 - \pi_{M_1}$. Then we would have:

$$\omega'_1 = \frac{1}{\pi_{M_1}}$$
$$\omega'_2 = \frac{1}{\pi_{M_2}} \tag{49}$$

Once again, we could see that as long as $M \geq 2$, we will have $\omega'_1 \geq \chi'_1$ and $\omega'_2 \geq \chi'_2$. We can thus rewrite the final set of sufficient conditions in the following manner:

$$\begin{gathered}
(1) \quad q(S_j, S_w, S_r, S_t) \leq 0 \\
(2) \quad g(S_j, S_w) \leq 0 \\
(3) \quad |g(S_j, S_w)| \geq \\
max\{\lambda_2 \log \omega'_1 + \frac{\lambda_1}{P(G_1)} \log{(\omega'_1\omega'_2)}, \\
\lambda_2 \log \omega'_2 + \frac{\lambda_1}{P(G_2)} \log{(\omega'_2\omega'_1)}\} \\
, \omega'_1 = \frac{1}{\pi_{M_1}}, \omega'_2 = \frac{1}{\pi_{M_2}}
\end{gathered}$$

Now that we have found the set of sufficient conditions for $F(S_j)$ in both cases of Eq.(14) and Eq.(15), we can find the overall set of $F(S_j)$ in all cases as:

$$\begin{gathered}
(1) \quad q(S_j, S_w, S_r, S_t) \leq 0 \\
(2) \quad g(S_j, S_w) \leq 0 \\
(3) \quad |g(S_j, S_w)| \geq \\
max\{\lambda_2 \log \omega'_1 + \frac{\lambda_1}{P(G_1)} \log{(\omega'_1\omega'_2)}, \\
\lambda_2 \log \omega'_2 + \frac{\lambda_1}{P(G_2)} \log{(\omega'_2\omega'_1)}\} \\
, \omega'_1 = \frac{1}{\pi_{M_1}}, \omega'_2 = \frac{1}{\pi_{M_2}}
\end{gathered}$$

□